Structure, Vol. 11, 845–853, July, 2003, ©2003 Elsevier Science Ltd. All rights reserved. DOI 10.1016/S0969-2126(03)00125-4

Molecular Roots of Degenerate Specificity in Syntenin's PDZ2 Domain: Reassessment of the PDZ Recognition Paradigm

Beom Sik Kang, David R. Cooper, Yancho Devedjiev, Urszula Derewenda, and Zygmunt S. Derewenda*
Department of Molecular Physiology and Biological Physics
The Cancer Center
University of Virginia
Charlottesville, Virginia 22908

Summary

Crystal structures of the PDZ2 domain of the scaffolding protein syntenin, both unbound and in complexes with peptides derived from C termini of IL5 receptor (α chain) and syndecan, reveal the molecular roots of syntenin's degenerate specificity. Three distinct binding sites (S_0, S_{-1}, and S_{-2}), with affinities for hydrophobic side chains, function in a combinatorial way: S_{-1} and S_{-2} act together to bind syndecan, while S_0 and S_{-1} are involved in the binding of IL5Rα. Neither mode of interaction is consistent with the prior classification scheme, which defined the IL5Rα interaction as class I (-S/T-X-ϕ) and the syndecan interaction as class II (-ϕ-X-ϕ). These results, in conjunction with other emerging structural data on PDZ domains, call for a revision of their classification and of the existing model of their mechanism.

Introduction

PDZ domains (postsynaptic density protein, disc large, and *zonula occludens*) occur within numerous multidomain cytosolic proteins and mediate their binding to receptors and channels, thereby serving as a membrane-associated scaffold for the assembly of signaling complexes (Harris and Lim, 2001; Hung and Sheng, 2002). Over 440 domains of this type have been identified so far in the human genome, and they are also abundant in other organisms (Sheng and Sala, 2001). PDZ domains are structurally conserved modules about 90 amino acids in size. The majority are believed to function by binding the C-terminal tail of the target protein in a structurally conserved groove between the β2 strand and the α2 helix (Doyle et al., 1996). The terminal carboxylate of the target is anchored via hydrogen bonds from three main chain amides within a conserved glycine-rich loop, a fingerprint of the PDZ fold. Early data derived from crystallographic and NMR studies suggested a general model of sequence pattern recognition, in which the peptide is bound in an extended conformation so that two side chains, P_0 and P_{-2}, point into the groove of the PDZ domain and account for specificity (P_0 denotes the C-terminal residue of the bound peptide and P_{-n} denotes the *n*th amino acid upstream of it; S_{-n} denotes the corresponding binding pocket of the PDZ domain). Those domains that are grouped together as class I bind Ser or Thr in P_{-2} and a hydrophobic residue in P_0, so that the target sequence motif is -S/T-X-ϕ (ϕ represents hydrophobic residues and Ψ represents aromatic residues). Class II domains bind another hydrophobic residue at P_{-2} (-ϕ-X-ϕ), while a negatively charged residue at P_{-2} defines class III interactions (-D/E-X-ϕ). This simple model is unable to explain an increasing number of PDZ-mediated interactions that do not conform to this canonical type of recognition. To account for them, new classes of PDZ domains are being proposed to extend the model. For example, PDZ1 of Mint1 has been termed "novel class III" (-E/D-X-W-C/S) (Maximov et al., 1999), and PDZ3 of hINADL has been placed in "class IV" (-X-Ψ-D/E) (Vaccaro and Dente, 2002). These new classes of PDZ domains recognize P_{-1} instead P_{-2}. To further complicate the issue, some PDZ domains recognize more than one class of the C-terminal sequence motif. CIPP PDZ3 binds neurexin (class II) and the NMDA receptor (class I) (Kurschner et al., 1998), and the third PDZ domain of hINADL binds the sequences -Ψ-D-ϕ (class II) and -X-Ψ-D sequence (class IV) (Vaccaro and Dente, 2002), while MINT1 PDZ1, hINADL PDZ5, and Par6 PDZ domains bind ligands with sequences -D-H-W-C (novel class III) and -E-Y-Y-V (class II) (Bezprozvanny and Maximov, 2001). The erbin PDZ domain binds the receptor ErbB2 (class II) and LET-23 peptide (class I) (Borg et al., 2000). While dual specificity is not rare in PDZ binding, there is no general model accounting for it.

Syntenin, first identified as a syndecan binding protein, contains a tandem of PDZ domains, which demonstrate degenerate specificity (Grootjans et al., 1997). In addition to syndecan, there are currently at least 10 binding partners reported for syntenin, including class I proteins such as interleukin 5 receptor α subunit (IL5Rα) (-D-S-V-F) (Geijsen et al., 2001), neuroglian (-Y-S-L-A) (Koroll et al., 2001), proTGF-α (-E-T-V-V) (Fernandez-Larrea et al., 1999), and neurofascin (-Y-S-L-A) (Koroll et al., 2001); class II molecules such as syndecan (-E-F-Y-A), ephrin B (-Y-Y-K-V) (Lin et al., 1999; Torres et al., 1998), Eph A7 (-G-I-Q-V) (Torres et al., 1998), PTP-η (-G-Y-I-A) (Iuliano et al., 2001), and neurexin I (-E-Y-Y-V) (Grootjans et al., 2000); and the class III protein merlin (-F-E-E-L) (Jannatipour et al., 2001). In principle, such diversity of interactions could be caused by degenerate specificity or alternatively by cooperative effects of two PDZ domains. We recently showed that syntenin's two PDZ domains show degenerate and noncooperative binding (Kang et al., 2003). The second PDZ domain (PDZ2) binds IL5Rα (class I) and syndecan-4 (class II) peptides, in spite of dramatically dissimilar sequences, with the dissociation constants of 1.9 μM and 2.3 μM, respectively. Mutational studies also show that PDZ2 has binding capacity for both class I and class II peptides (Grootjans et al., 2000; Koroll et al., 2001). In order to elucidate the molecular basis for the dual specificity of the PDZ2 domain of syntenin, we determined the crystal structures of the PDZ2 domain alone and in com-

*Correspondence: zsd4n@virginia.edu

Table 1. Crystallographic Data

Data Set	PDZ2		PDZ2-Syndecan-4	PDZ2-IL5Rα
Experimental Data				
Wavelength (λ)	1.54178	0.97946	0.97946	0.97946
Space group	$P2_1$	$P2_1$	C2	$C222_1$
Unit cell parameters (Å, °)				
a	25.27	25.29	58.34	53.72
b	42.54	42.57	54.44	55.98
c	31.06	31.04	50.22	51.09
β	108.8	108.7	98.7	90.0
Resolution (Å)	25.0–1.60 (1.66–1.60)[a]	13.0–1.10 (1.14–1.10)	50.0–1.85 (1.92–1.85)	20.0–1.25 (1.29–1.25)
Total reflections	21,360	58,393	46,421	155,256
Unique reflections	6,600 (186)	18,559 (275)	12,987 (1,108)	20,850 (1,611)
Completeness (%)	79.8 (22.9)	73.3 (11.0)[b]	97.4 (83.7)	95.7 (75.0)
R_{sym} (%)[c]	5.8 (17.0)	4.2 (17.0)	4.9 (41.3)	5.1 (49.6)
Average I/σ (I)	22.7 (3.62)	28.5 (3.46)	25.8 (2.86)	33.5 (2.15)
Refinement Details				
Resolution (Å)	21.27–1.60	12.16–1.24	49.39–1.85	19.39–1.35
Reflections (working)	6,290	15,685	12,001	16,171
Reflections (test)	307	851	985	848
R_{work} (%)[d]	11.9	11.3	17.5	17.6
R_{free} (%)[d]	16.6	15.3	22.6	21.2
Number of waters	181	173	145	144
Rms deviation from ideal geometry				
Bonds (Å)	0.015	0.013	0.012	0.016
Angles (°)	1.48	1.63	1.82	2.13
Average B factor (Å^2)				
Main chain	11.42	8.32	16.01	17.19
Side chain	12.69	10.15	20.50	21.50
Waters	26.41	22.21	39.86	39.93

[a] The numbers in parentheses describe the relevant value for the last resolution shell.
[b] Completeness at resolution 12.16–1.24 (1.30–1.24) used for refinement is 93.1% (73.4%)
[c] $R_{sym} = \Sigma||I_i - <I>|/\Sigma I$ where I_i is the intensity of the *i*th observation and $<I>$ is the mean intensity of the reflections.
[d] $R_{work} = \Sigma||F_{obs}| - |F_{calc}||/\Sigma |F_{obs}|$, crystallographic R factor, and $R_{free} = \Sigma||F_{obs}| - |F_{calc}||/\Sigma|F_{obs}|$ when all reflections belong to a test set of randomly selected data.

plexes with an IL5Rα C-terminal peptide (ETLEDSVF) and a syndecan-4 peptide (TNEFYA). The structures were refined to 1.24 Å, 1.35 Å, and 1.85 Å resolution, respectively. These structures show how syntenin's PDZ2 can accommodate different peptides and call for a revision of the established paradigm of PDZ domain classification.

Results and Discussion

Syndecan Binding Involves Interaction with Tyr_{-1}

The crystal structure of the PDZ2 domain with a bound syndecan-4 C-terminal hexapeptide was refined at 1.85 Å resolution to a crystallographic R value of 17.5% and R_{free} of 22.6% (Table 1). The structure contains a noncrystallographic dimer of PDZ2-peptide complexes in the asymmetric unit. In both complexes, the structures of the bound syndecan-4 peptides were identical within experimental error, with the average isotropic temperature factor (B factor) of 30.9 Å^2 and 31.4 Å^2, respectively. In general, the interaction between the PDZ2 and the peptide conforms to the classical model of a strand insertion between the β2 strand and α2 helix of the PDZ domain (Figure 1A). The terminal carboxylate of the peptide accepts three hydrogen bonds from the amide nitrogens of Val209, Gly210, and Phe211. There is an additional indirect interaction with the carbonyl oxygen of Gly207 through an ordered water molecule. The main chain amide of the C-terminal residue donates a hydrogen bond to the carbonyl oxygen of Phe211 in the β2 strand. The carbonyl oxygen of Phe (P_{-2}) interacts with the amide group of Phe213, while its amide donates a hydrogen bond to the carbonyl oxygen of the same residue. The C-terminal Ala (P_0) and Phe (P_{-2}) of the syndecan peptide interact with PDZ2 in agreement with the canonical model of class II, as exemplified by the structure of hCASK (Daniels et al., 1998). However, the methyl group of Ala (P_0) is much smaller than the size of the hydrophobic pocket at S_0, which is formed mostly by Val209, Phe211, Phe213, and Leu258 (Figure 1B). In contrast, the benzene ring of Phe (P_{-2}) fits well in the corresponding S_{-2} pocket formed by Phe213, Asp251, and Ala255.

Interestingly, there is an additional interaction involving Tyr (P_{-1}), which is lodged into the S_{-1} pocket cushioned by His208, Ileu212, and Val222. The aromatic ring is involved in an off-center stacking interaction with His208. Other syntenin PDZ2 binding proteins, neurexin I (class II), and neuroglian (class I) also have Tyr in the P_{-1} position (Grootjans et al., 2000; Koroll et al., 2001). The equal importance of the Phe (P_{-2}) and Tyr (P_{-1}) interactions for recognition by syntenin's PDZ2 is underscored by studies showing that mutations to Ala of either of the two residues abolish binding to syntenin (Groot-

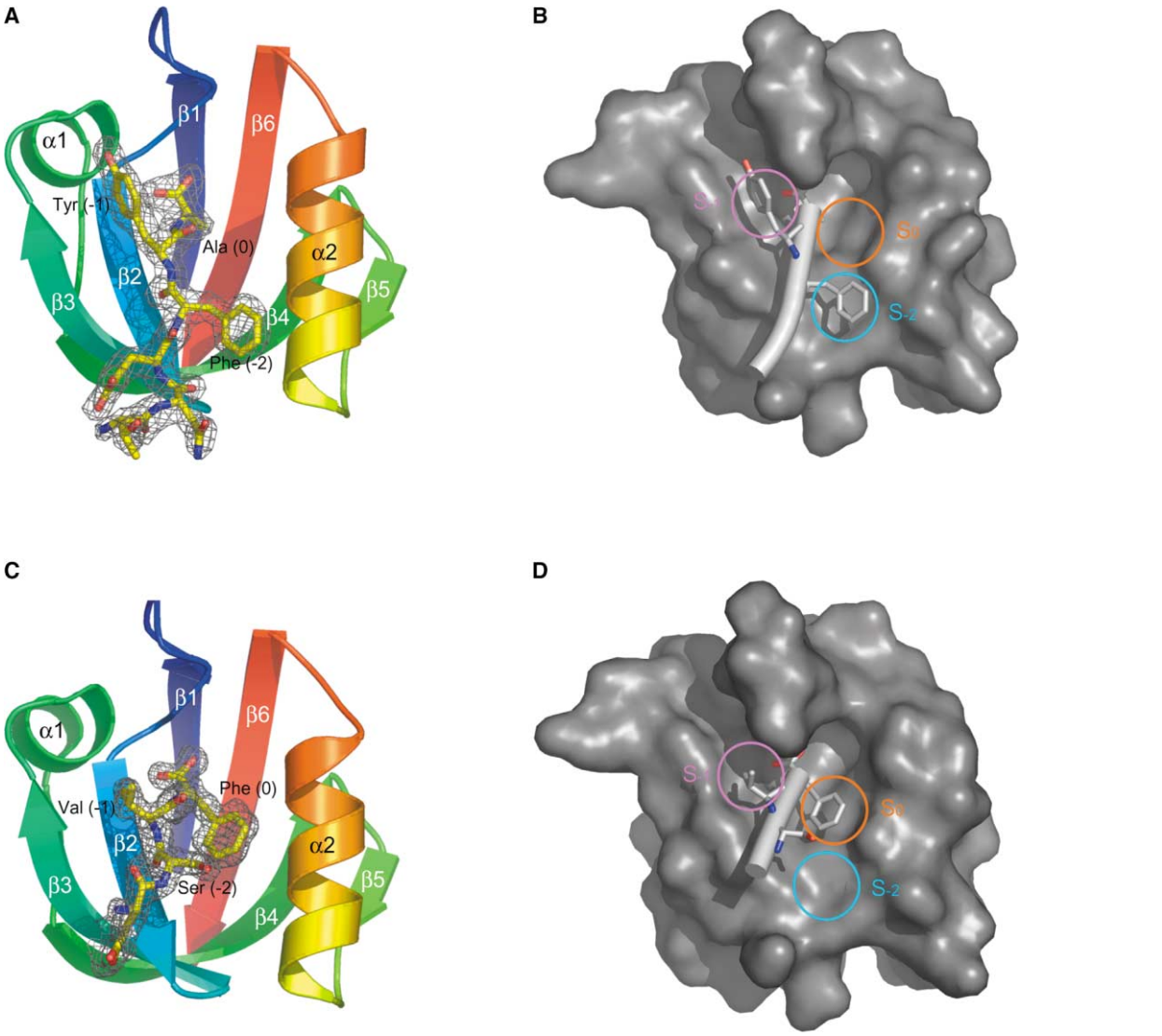

Figure 1. Comparison of Syntenin PDZ2 Structures Binding Syndecan-4 Peptide and Interleukin 5 Receptor α Subunit

(A) Ribbon diagram of the syntenin PDZ2 bound to the syndecan-4 peptide (TNEFYA). A $2mF_0 - DF_c$ electron density map calculated at 1.85 Å resolution and contoured at 1.0σ is shown around the ligand.

(B) Molecular surfaces of syntenin PDZ2 showing three hydrophobic binding pockets and the syndecan-4 peptide. The three binding pockets are circled. The three C-terminal residues are shown in the Cα trace cartoon of the peptide. The side chains of tyrosine (−1) and phenylalanine (−2) occupy the two pockets S_{-1} and S_{-2}, while alanine (0) only occupies a portion of S_0.

(C) Ribbon diagram of the syntenin PDZ2 bound to the interleukin 5 receptor α subunit peptide (ETLEDSVF). A $2mF_0 - DF_c$ electron density map calculated at 1.35 Å resolution and contoured at 1.0σ is shown around the ligand.

(D) Molecular surfaces of syntenin PDZ2 showing three hydrophobic binding pockets and the interleukin 5 receptor α subunit peptide. The three binding pockets are circled. The three C-terminal residues are shown in the Cα trace cartoon of the peptide. The side chains of phenylalanine (0) and valine (−1) of the peptide are located in pockets S_0 and S_{-1}, while that of serine (−2) is out of pocket S_{-2}. The same orientation is used for (A) and (B) or (C) and (D). Figures were made using MOLSCRIPT (Kraulis, 1991) and Pymol (DeLano Scientific). Strand β1 (197–203); β2 (210–214); β3 (216–222); β4 (234–242); β5 (244–246); β6 (A263–270); Helix α1 (225–231); α2 (250–260).

jans et al., 1997). Thus, the interaction of PDZ2 with syndecan depends primarily on the side chains of residues in P_{-1} and P_{-2} rather than P_0 and P_{-2}, as the classical model implies.

The Interaction of IL5Rα with Syntenin Does Not Involve Ser_{-2}

The interaction of syntenin with IL5Rα was originally reported as class I, because the C-terminal sequence of the receptor has Ser in the P_{-2} position and Phe in the P_0 position. However, the sequences of syntenin PDZ domains do not resemble a typical class I domain. In particular, there is a notable absence of a histidine at the beginning of helix α2, which normally hydrogen bonds to the hydroxyl of Ser or Thr (P_{-2}). In an effort to characterize the details of the IL5Rα interactions with syntenin, we crystallized and solved the structure of the PDZ2 domain with an octapeptide derived from the

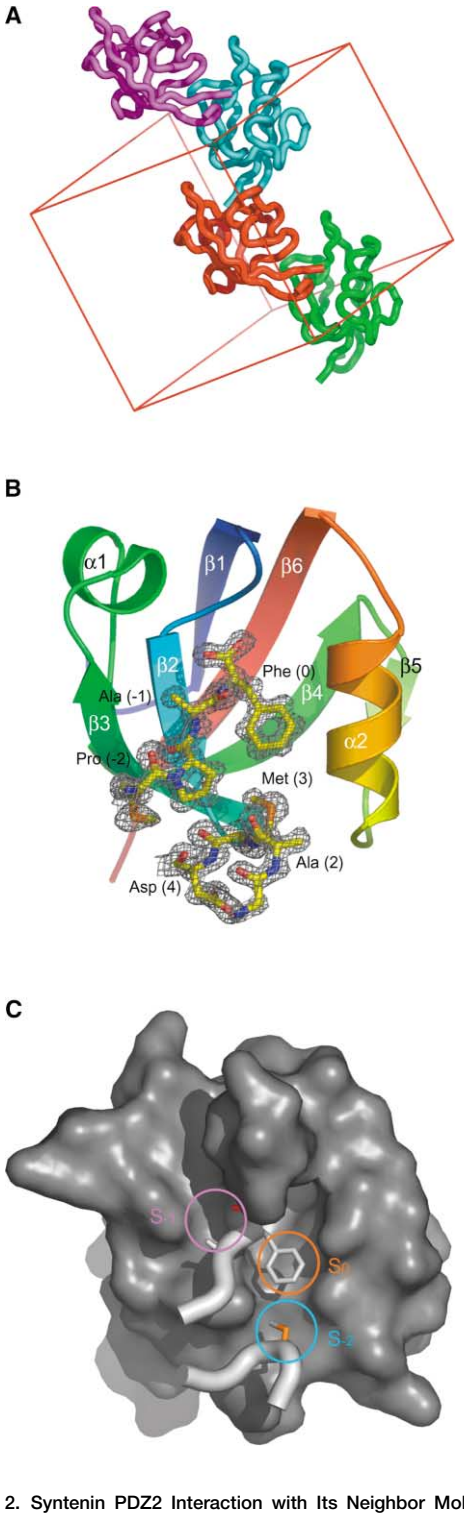

Figure 2. Syntenin PDZ2 Interaction with Its Neighbor Molecular Shows Novel PDZ Binding

(A) Crystal packing of syntenin PDZ2. Each C terminus serves as a ligand for a neighboring PDZ2 molecule.

(B) Ribbon diagram of syntenin PDZ2 bound to the C-terminal and internal sequences of its neighbor molecule. A $2mF_0 - DF_C$ electron density map calculated at 1.24 Å resolution and contoured at 1.0σ around the residues bound in the PDZ domain.

(C) Molecular surfaces of syntenin PDZ2 showing three hydrophobic binding pockets and the residues of neighboring molecule binding at the pockets. The three binding pockets are circled. The three residues are shown in the Cα trace cartoon of bound residues. The side chains of phenylalanine (0) and alanine (−1) of the C terminus reside in pockets S_0 and S_{-1}, while that of methionine (3) is in pocket S_{-2}. The same orientation is used for (B) and (C). Figures were made using MOLSCRIPT (Kraulis, 1991) and Pymol (DeLano Scientific).

C-terminal sequence of IL5Rα. The crystals of the complex allowed for X-ray data collection to a resolution of 1.35 Å (Table 1). The atomic model, refined to a crystallographic R value of 17.6% (R_{free} 21.2%), shows how the C-terminal carboxylate group of Phe (P_0) of the peptide is bound in a way analogous to that seen in the syndecan complex, while its benzene ring fills the S_0 pocket (Figures 1C and 1D). However, Ser (P_{-2}) does not directly interact with the PDZ2 domain as suggested by the classical model. The side chain of Val (P_{-1}) fits into the hydrophobic S_{-1} pocket, and the carbonyl oxygen of Ser (P_{-2}) is hydrogen bonded through a water molecule to the Ile212 main chain nitrogen. There are no further interactions, and the electron density for the peptide fades beyond P_{-4}. Except for the interaction of Phe (P_0), the peptide's backbone does not fully occupy the binding groove as is seen in other complexes, leaving the S_{-2} site empty. Similar interaction of these three C-terminal residues was also found in the crystal structure of the syntenin PDZ tandem-IL5Rα peptide complex (data not shown, our PDB entry 1OBZ). It has been shown by mutational studies that with the exception of the C-terminal Phe, no other residue is vital for the interaction of IL5Rα with syntenin (Geijsen et al., 2001). This is in agreement with our results but does not support the classical class I recognition mechanism. In contrast, the structure reveals some similarities to the conformation observed in erbin-ErbB2 peptide complex (Birrane et al., 2003). Failure of Val (P_{-2}) to form a typical class I interaction with His at α2 helix causes the displacement of the peptide backbone away from the α helix. In the complex of erbin and phosphorylated ErbB2 peptide (EpYLGLDVPV) complex, no density is observed beyond P_{-5}, leaving its other binding pocket at the β2-β3 loop empty.

The Structure of Unbound PDZ2 Suggests Additional Recognition Mechanisms

PDZ domains show high affinity toward terminal carboxyl groups of peptides, and in the absence of target peptides, isolated PDZ domains will often bind their own C-terminal tails. The crystal structures of the hCASK PDZ, NHERF PDZ1, and GRIP1 PDZ6 domains show how the peptide binding grooves are occupied by the C-terminal tails of neighboring molecule, mimicking the recognition of the peptide ligand (Daniels et al., 1998; Im et al., 2003; Karthikeyan et al., 2001a). In a similar way, the crystal structure of the uncomplexed PDZ2 domain, refined at 1.24 Å resolution to a crystallographic R value of 11.3% (R_{free} 15.3%), shows an interesting interaction between adjacent molecules (Table 1 and Figure 2A). This interaction suggests that PDZ2 is capable of yet another mode of molecular recognition, in addition to the classic mode of strand insertion.

The PDZ2 construct used here terminates at residue Phe273, rather than Met270, as is the case with the

crystal structures of the two above described complexes. (Chronologically, this structure was done first, and the construct was then truncated to circumvent inter-PDZ2 interactions and inhibition of peptide binding.) One molecule of PDZ2 binds the C-terminal Phe of its crystallographic neighbor in a manner identical to that observed for the IL5Rα (Figures 2B and 2C). The preceding residues are out of the binding groove, and there is no interaction of P_{-2} at S_{-2}. Compared to either the syndecan or IL5Rα bound structures, the binding groove is not altered except for the side chain of Ile212, which rotates to accommodate the C terminus of the neighboring molecule. In addition to the interaction involving the C-terminal Phe, the N-terminal portion of the same molecule also interacts with the binding groove. This is possible because the C terminus and N terminus of PDZ2 are close to each other and form a structural epitope. The side chain of the third residue Met binds into the S_{-2} site, so that its methyl group is in a very close contact with the aromatic ring of Phe213. A salt bridge forms between carboxyl group of an Asp, which follows the Met, and amino group of Lys214 in β2 strand of the adjacent molecule. Thus, in this case, syntenin's PDZ2 domain shows affinity for a *structural* epitope, rather than a *sequence* motif. It is very probable that this mode of recognition also occurs in nature, and that the two binding sites (S_0 and S_{-2}) may bind amino acids, which need not be adjacent within a short peptide.

Molecular Basis of Recognition by PDZ Domains

Our results, in conjunction with existing literature, call into question the utility of the current model of protein recognition by PDZ domains and of the rigid classification of PDZ domains based on the identity of P_0 and P_{-2} residues of a target peptide (Figure 3A). As shown here, syntenin PDZ2 has three distinct binding pockets (S_0, S_{-1}, and S_{-2}), and the interaction of the P_{-1} residue at the S_{-1} site is as important as the canonical interactions at the S_0 and S_{-2} sites. Therefore, P_{-1} interaction should be included as a general feature of PDZ interaction (Figure 3B). The importance of the P_{-1} binding is also apparent from studies of other PDZ domains. Both LIN-2 and p55 PDZ domains bind peptides where all three terminal residues are hydrophobic, including aromatic side chains at both P_{-1} and P_{-2} (Songyang et al., 1997). By phage-displayed peptide library screening, PDZ2 of MAGI3 selects Trp as P_{-1} and the side chain of this Trp is critical for high-affinity binding (Fuh et al., 2000). Model studies suggest that the side chain at P_{-1} position reaches across the β2 strand and makes specific contacts with side chains in the β3 strand. The binding specificity studies of hINADL reveal that PDZ1, 2, 3, and 4 belong to class II while PDZ5, 6, and 7 are class I PDZ domains (Vaccaro et al., 2001). However, except for PDZ7, all domains in hINADL have some selectivity for P_{-1}, and the site-directed mutagenesis of the residues in β3 of hINADL PDZ7 alters the selectivity for P_{-1}. The atomic model of the NHERF PDZ domain complexed with the -Q-D-T-R-L target sequence also reveals recognition mediated by the P_{-1} residue: the Arg side chain in this position interacts intimately with Asn22 and Glu43 of the PDZ (Karthikeyan et al., 2001b); these residues, lining the S_{-1} pocket, are equivalent to His208 and Val222 of syntenin PDZ2.

Another example is the erbin PDZ domain, normally defined as a class I because of the His at the beginning of the α2 helix and because of the target motif -S/T-X-ϕ. Interestingly, this domain does not bind ErbB4 (-N-T-V-V) (Borg et al., 2000), but interacts with the C termini of δ-catenin, p0071, and ARVCF, which all share the sequence -D-S-W-V (Laura et al., 2002). Furthermore, Trp was exclusively selected for P_{-1} for the erbin PDZ domain by phage display peptide library screening. The NMR solution structure of the erbin PDZ domain with the phage-optimized peptide (AcTGWETWV) reveals canonical interactions of P_0 and P_{-2} residues, as well as an additional interaction involving Trp (P_{-1}), whose side chain reaches across strand β2 and inserts between the side chain of Arg49 and Gln51 at the end of β3 strand (Skelton et al., 2003). Clearly, a more accurate description of the target motif for erbin would be -S/T-W-ϕ, highlighting the preference at P_{-1} for Trp.

In syntenin's PDZ2, all three binding pockets (S_0, S_{-1}, and S_{-2}) have an apparent affinity for hydrophobic residues (-ϕ-ϕ-ϕ). These pockets function in a combinatorial way to bind peptides from different targets. This interplay of the three sites appears to endow PDZ2 with the ability to bind diverse but specific sequence motifs. The complex of syntenin's PDZ2 domain with syndecan-4 peptide, previously classified as class II interaction, involves the recognition of two penultimate residues, Tyr (P_{-1}) and Phe (P_{-2}), but not of the side chain of P_0 (-ϕ-ϕ-X) (Figures 1B and 3C). On the other hand, the interaction with the IL5Rα peptide does not involve the P_{-2} side chain hydroxyl, unlike other PDZ domains that interact with class I peptides (Figure 1D). Therefore, defining the IL5Rα peptide as class I partner for syntenin is questionable. Our data suggest that syntenin's PDZ2 interactions with other so-called class I peptides may also involve binding of P_0 and P_{-1} (-X-ϕ-ϕ) instead of P_0 and P_{-2} as the classical model requires (Figure 3D). All of syntenin's partners known to bind to PDZ2 have hydrophobic residues in P_{-1}.

Our combinatorial model incorporates the classical P_0 and P_{-2} interactions (-ϕ-X-ϕ) but accounts for all three S sites (-ϕ-ϕ-ϕ) (Figures 3A and 3B). We expect that the former mode expains syntenin's interaction with ephrin B (-Y-Y-K-V), while the latter may apply to neurexin (-E-Y-Y-V). This model could also explain the dual specificity observed for other PDZ domains. For example, the CIPP PDZ3 domain binds NMDA receptors (class I) and neurexin (class II). Although the former target has a sequence -E-S-D/E-V, the CIPP PDZ3 domain does not have a His at α2 and it does not bind a related peptide from neuroligin2 (-T-T-R-V) (Kurschner et al., 1998), indicating that the P_{-2} position is not critical, in contrast to canonical class I interaction. The SMART database places this particular PDZ domain in a group that binds a motif -Ψ-D/E-ϕ (Bezprozvanny and Maximov, 2001). Applying the combinatorial three-pocket model helps understand how it can bind NMDA receptors using P_{-1} and P_0 (-S-D/E-V) and neurexin using P_{-2} and P_0 (-Y-Y-V).

Aside from explicit examples of the involvement of the P_{-1} residue—unaccounted for in the classical

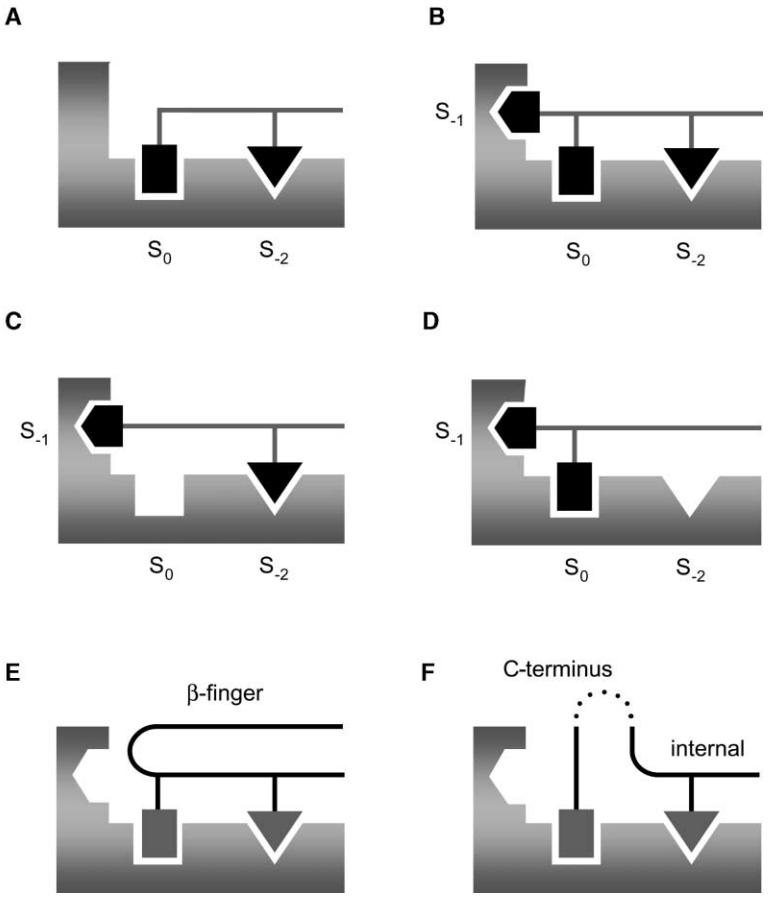

Figure 3. Schematics of PDZ Interactions

(A) Canonical PDZ binding of C-terminal sequence depends on the residues P_0 and P_{-2} binding pockets S_0 and S_{-2}.

(B) All three C-terminal residues are involved in the interaction to PDZ binding groove.

(C) The C-terminal binding depends on the binding at S_{-1} and S_{-2} of PDZ as seen in PDZ2-syndecan-4 peptide complex.

(D) The C-terminal binding depends on the binding at S_0 and S_{-1} as seen in syntenin PDZ2-interleukin 5 receptor α subunit peptide complex.

(E) Syntrophin PDZ interaction by the residues from β-finger conformation of nNOS.

(F) Interaction of internal residue at pocket S_{-2} while C-terminal residues binds at S_0 as seen in syntenin PDZ2-PDZ2 interaction.

model—there are examples of the P_{-2} residue not being involved, as in the interaction of syntenin's PDZ2 with IL5Rα. The so-called class IV or class III interactions, targeting sequences -X-Ψ-D/E or -X-W-C, do not show specificity for P_{-2} (Bezprozvanny and Maximov, 2001; Maximov et al., 1999). It is not surprising that the degenerate specificity is common among these PDZ domains. Mint1 PDZ1, which has dual specificity for sequences -E/D-X-W-C/S and -E-Y-Y-V, could bind both types of sequences by interacting with P_{-1}/P_0 and P_{-2}/P_{-1} pairs, respectively. The ability of the fifth PDZ domain of hINADL (Vaccaro et al., 2001) and PAR6 PDZ domain to bind the sequences -D-H-W-C and -E-Y-Y-V may be rationalized in a manner similar to the Mint1 PDZ1 interaction. However, there may be additional interaction involving the conserved P_{-3}, residue, D/E.

We believe that all three S sites described here are key determinants in the PDZ complex recognition pattern, and any general model should include all three. The binding modes of known PDZ domains, according to our combinatorial model depicted in Figure 3, are summarized in Table 2.

There is also accumulating evidence that residues upstream of the C-terminal tripeptide of the target may also be involved in the recognition process, putting the classical model in even greater peril (Laura et al., 2002; Songyang et al., 1997; Vaccaro et al., 2001). Although we found no interaction with P_{-3} in the syntenin PDZ2 structures, the specificity for P_{-3} is often observed. Some PDZ domains, which have long β2 strand or β2-β3 loop, have further interaction to its target peptide at this region (Birrane et al., 2003; Kozlov et al., 2002; Walma et al., 2002). We have shown that the PDZ1 domain of syntenin, which also has a long β2-β3 loop, recognizes residues upstream of the terminal hexapeptide of merlin, as exemplified by a significantly higher affinity for octathan hexapeptide (Kang et al., 2003). However, syntenin PDZ2 has a short β2-β3 loop like PSD-95 PDZ1 and NHERF PDZ1, and these domains do not interact with the bound peptides at this loop (Doyle et al., 1996; Karthikeyan et al., 2001b, 2002). Thus, the interaction including P_{-3} appears to be optional, and it could enhance the binding in the absence of, or in addition to, strong binding at C-terminal three residues.

Finally, the recent erbin structures show how a binding pocket can be targeted by a residue that does not occupy the expected sequence position. The crystal structure with ErbB2 peptide shows that Tyr_{-7} of the peptide binds at a site within the β2-β3 loop (Birrane et al., 2003). Interestingly, in the structure with the phage peptide (AcTGWETWV), Trp_{-4} interacts at the same binding site (Skelton et al., 2003). This implies that a structural epitope is more important than the sequence for PDZ-peptide interaction. As shown by our structure of syntenin's PDZ2 and its interaction with its neighboring molecule, the residues far from the C terminus in sequence can be involved in binding to the S_{-2} pocket by forming a contiguous structural epitope with the C terminus. There is also the exceptional example of the recognition of nNOS β-finger, not C terminus, by syntrophin (Figure

Table 2. Proposed Binding Modes of PDZ Domains by the Combinatorial Model

PDZ Domains	The C-Terminal Sequence Motifs of Representive Ligands	Binding Mode[a]	References
Class I (-S/T-X-ϕ)[b]			
PSD-95 PDZ3	-T-S-V	A[c]	Doyle et al., 1996
NHERF PDZ1	-T-R-L, -S-L-L, -S-F-L, -E-Q-L	A[c], B[c]	Karthikeyan et al., 2001a, 2001b, 2002
Erbin	-T-W-V, -V-P-V, -S-W-V	A[c], B[c]	Borg et al., 2000; Laura et al., 2002; Skelton et al., 2003
hPTP1E	-S-A-V	A[c]	Kozlov et al., 2002
Syntrophin	-S/T-X-V, S-L-V, T-T-F[d]	A[c], B[c]	Hillier et al., 1999; Schultz et al., 1998
MAGI3 PDZ2	-S/T-W-V	B	Fuh et al., 2000
CIPP PDZ3	-S-D/E-V, -Y-Y-V	A, C	Bezprozvanny and Maximov, 2001; Kurschner et al., 1998
Class II (-ϕ-X-ϕ)			
hCASK	-R-E-F, -F-Y-A	A[c], C	Daniels et al., 1998
Syntenin	-F-Y-A, -S-V-F, -Y-Y-V, -Y-K-V	A, B, C[c], D[c]	Geijsen et al., 2001; Grootjans et al., 1997, 2000; Torres et al., 1998; this study
GRIP PDZ6	-Y-S-C	B[c]	Im et al., 2003
P55	-Y-F-I, -F-X-X, -Y-Y-F	B, C	Songyang et al., 1997
LIN-2	-F-F-V/F/A	B, C	Songyang et al., 1997
Class III (-D/E-X-ϕ)			
nNOS	-D-S-V	A[d]	Tochio et al., 1999
Class IV (-X-Ψ-D/E, -X-W-C/S)			
hINADL PDZ3	-W-D-V, -Y-D-W, S-W-E, -S-Y-E	B, D	Vaccaro et al., 2001
Mint1 PDZ1	-D-W-C, -H-W-C, -Y-Y-V	C, D	Bezprozvanny and Maximov, 2001; Maximov et al., 1999
PAR6	-H-W-C, -Y-Y-V	C, D	Bezprozvanny and Maximov, 2001
hINADL PDZ5	-H-W-C, -Y-Y-V, -V-F-V	C, D	Bezprozvanny and Maximov, 2001; Vaccaro et al., 2001

[a] A, canonical P_0 and P_{-2} binding; B, P_0, P_{-1}, and P_{-2} binding; C, P_{-1} and P_{-2} binding; D, P_0 and P_{-1} binding, as shown in Figure 3.
[b] X denotes any amino acid; ϕ denotes a hydrophobic residue; Ψ denotes an aromatic residue.
[c] Structure was solved by X-ray or NMR.
[d] β-finger binding.

3E; Hillier et al., 1999), but the mode we suggest implies involvement of internal sequences *in addition to* the interaction of the C terminus in PDZ interaction (Figure 3F).

One of the benefits of PDZ classification is the potential ability to predict the binding partners for a given PDZ domain. However, even a correctly predicted single C-terminal sequence motif may not be enough to determine the binding capacity of any given PDZ domain. For successful prediction of multiple binding partners for PDZ domains, it would be better to characterize the PDZ domain by the specificity of its binding pockets, especially the three pockets for the C-terminal residues, and account for the likely combinations.

Biological Implications

Protein-protein interactions are pivotal to cell signaling events. The PDZ domain is the most ubiquitous protein-protein interaction module found in the human genome, with nearly 500 copies. Numerous multidomain cytosolic proteins contain PDZ domains and bind to receptors and channels, thereby serving as a membrane-associated scaffold for the assembly of signaling complexes. Syntenin is a ubiquitous protein containing two PDZ domains and is involved in protein targeting and multiprotein assembly. Notably, it is overexpressed in gastric and breast cancer cell lines, suggesting that its function contributes to cytoskeleton regulation and cell migration. Syntenin binds biologically important receptors such as IL5R and syndecan. We report the crystal structures of complexes of the second PDZ domain of syntenin, residues 113–270, with the C-terminal peptides of IL5Rα and syndecan, solved at 1.35 Å and 1.85 Å resolution, respectively. These structures reveal how one PDZ domain interacts with different C-terminal sequences of binding targets. Syndecan binds syntenin mainly by its C-terminal P_{-1} and P_{-2} residues, and IL5Rα interacts through its C-terminal P_0 and P_{-1} residues. These results not only extend the knowledge of PDZ-ligand recognition of specific targets but also explain the general scheme underlying degenerate specificity. Furthermore, the mode of syntenin PDZ2 interaction with its neighbor molecule in a crystal of the unbound PDZ2 domain (1.24 Å resolution) suggests the importance of a structural epitope for PDZ interactions rather than a sequence motif. Based on our results and the results published elsewhere, we propose the combinatorial model that generalizes the PDZ-ligand interactions. The new model is likely to predict the possible binding of biologically important target molecules more accurately than the current model.

Experimental Procedures

Protein Expression and Purification

A syntenin clone was obtained from American Tissue and Culture Collection (ATCC 72537). The DNA fragment coding for PDZ2 (197–273) was amplified by PCR and cloned into a GST-fusion expression vector containing the TEV (tobacco etch virus) protease cleavage

site (Sheffield et al., 1999). This construct was used for expression of protein samples for the crystallization of uncomplexed PDZ2. To obtain a shorter version of PDZ2 (197–270), a stop codon (TGA) was introduced after Met270 by the QuikChange® method (Stratagene). This step was necessary to prevent intermolecular interactions between the PDZ2 domains, in which one molecule bound another via the terminal Phe (see text). Both versions of the PDZ2 domain were expressed in *E. coli* BL21 strain (Stratagene) and purified using glutathione-Sepharose 4B column (Amersham Pharmacia Biotech). The eluted recombinant protein was subjected to a HiPrep 26/10 desalting column (Amersham Pharmacia Biotech) equilibrated with 50 mM Tris-HCl (pH 7.5), 150 mM NaCl (buffer A) to remove glutathione. After complete digestion with rTEV protease (Life Technologies) at 10°C in the presence of 0.5 mM EDTA and 1 mM DTT, the protein solution was passed again through a glutathione Sepharose 4B column and the flow-through was concentrated and loaded on a Superdex G75 column (Amersham Pharmacia Biotech) equilibrated with buffer A. The protein fractions were collected and concentrated. All the purification steps, except the rTEV digestion, were performed at 4°C. The purified PDZ2 domain contains an additional pentapeptide (GAMDP) at the N terminus due to the cloning procedure.

Crystallization

Initial search for crystals of PDZ2 (197–273) was carried out with Crystal Screen 1 (Hampton Research, Inc.) using the hanging drop vapor-diffusion technique at 294K. The best crystals of PDZ2 were obtained with 8 mg/ml protein concentration at 0.1 M HEPES (pH 7.0), 34% PEG4000 using the sitting drop vapor-diffusion method with microseeding. For the crystallization of complexes of short PDZ2 (197–270) with peptides, we used Additive screen I (Hampton Research, Inc.) for additional screening. The complex of PDZ2 and syndecan-4 peptide was crystallized from 0.1 M HEPES (pH 6.8), 1.6 M ammonium sulfate, 20 mM $CoCl_2$, and 0.2 M $MgSO_4$, using 1:2 molar mixtures of protein and peptide. The best crystal of PDZ2 with the IL5Rα peptide was obtained from 0.1 M HEPES (pH 7.0), 1.6 M ammonium sulfate, 20 mM $CoCl_2$, and 0.2 M $MgSO_4$ by microseeding. Synthesized octapeptide of IL5Rα (ETLEDSVF) and hexapeptide of syndecan-4 (TNEFYA) were purchased from Bio-Synthesis and UVA Biomolecular Research Facility, respectively.

Data Collection, Structure Determination, and Refinement

Crystals were frozen in liquid nitrogen. Those containing peptides were briefly soaked in the crystallization buffer containing 17.5% (v/v) glycerol and peptide before freezing. An initial data set was collected with an R-Axis IV detector and a Nonius FR591 generator equipped with Osmic confocal mirrors. Subsequent data were collected at beamline X9B at NSLS with a wavelength of 0.97946 Å under cryoconditions using an ADSC Quantum4 CCD. Data sets were processed and scaled using HKL2000 (Otwinowski and Minor, 1997). Crystallographic details including unit cells and data statistics are shown in Table 1. All structures were solved by the molecular replacement method using AMORE (Navaza, 1994) and the atomic models were refined with REFMAC5 (Murshudov et al., 1997) from the CCP4 suite (CCP4, 1994). The atomic coordinates of the PDZ2 domain derived from the structure of the PDZ tandem (entry 1N99 in the PDB) were used as initial model for molecular replacement of the home source data set of PDZ2. The PDZ2 structure refined to 1.60 Å was subsequently used as a model for molecular replacement of the other data sets collected using synchrotron radiation. Manual model building was performed in O (Jones et al., 1991). The final models agree well with known protein geometry. Details of refinement are given in Table 1. Compared to unbound PDZ2 in the PDZ tandem structure, the rms differences for Cα atoms of PDZ2 structures with syndecan, IL5Rα peptide, and alone but interacting with the neighboring molecule, are 0.39 Å, 0.44 Å, and 0.39 Å, respectively, indicating that the bound peptide causes no significant structural changes in the PDZ domain.

Acknowledgments

Supported by the DOD grant DAMD17-01-1-0720 (to Z.S.D.). We thank Mary Lewis for excellent technical assistance and Dr. Zbigniew Dauter (NCI/NSLS) for help with data collection.

Received: February 13, 2003
Revised: March 3, 2003
Accepted: March 27, 2003
Published: July 1, 2003

References

Bezprozvanny, I., and Maximov, A. (2001). Classification of PDZ domains. FEBS Lett. *509*, 457–462.

Birrane, G., Chung, J., and Ladias, J.A. (2003). Novel mode of ligand recognition by the erbin PDZ domain. J. Biol. Chem. *278*, 1399–1402.

Borg, J.P., Marchetto, S., Le Bivic, A., Ollendorff, V., Jaulin-Bastard, F., Saito, H., Fournier, E., Adelaide, J., Margolis, B., and Birnbaum, D. (2000). ERBIN: a basolateral PDZ protein that interacts with the mammalian ERBB2/HER2 receptor. Nat. Cell Biol. *2*, 407–414.

CCP4 (Collaborative Computational Project Number 4) (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D *50*, 760–763.

Daniels, D.L., Cohen, A.R., Anderson, J.M., and Brunger, A.T. (1998). Crystal structure of the hCASK PDZ domain reveals the structural basis of class II PDZ domain target recognition. Nat. Struct. Biol. *5*, 317–325.

Doyle, D.A., Lee, A., Lewis, J., Kim, E., Sheng, M., and MacKinnon, R. (1996). Crystal structures of a complexed and peptide-free membrane protein-binding domain: molecular basis of peptide recognition by PDZ. Cell *85*, 1067–1076.

Fernandez-Larrea, J., Merlos-Suarez, A., Urena, J.M., Baselga, J., and Arribas, J. (1999). A role for a PDZ protein in the early secretory pathway for the targeting of proTGF-α to the cell surface. Mol. Cell *3*, 423–433.

Fuh, G., Pisabarro, M.T., Li, Y., Quan, C., Lasky, L.A., and Sidhu, S.S. (2000). Analysis of PDZ domain-ligand interactions using carboxyl-terminal phage display. J. Biol. Chem. *275*, 21486–21491.

Geijsen, N., Uings, I.J., Pals, C., Armstrong, J., McKinnon, M., Raaijmakers, J.A., Lammers, J.W., Koenderman, L., and Coffer, P.J. (2001). Cytokine-specific transcriptional regulation through an IL-5Rα interacting protein. Science *293*, 1136–1138.

Grootjans, J.J., Zimmermann, P., Reekmans, G., Smets, A., Degeest, G., Durr, J., and David, G. (1997). Syntenin, a PDZ protein that binds syndecan cytoplasmic domains. Proc. Natl. Acad. Sci. USA *94*, 13683–13688.

Grootjans, J.J., Reekmans, G., Ceulemans, H., and David, G. (2000). Syntenin-syndecan binding requires syndecan-synteny and the co-operation of both PDZ domains of syntenin. J. Biol. Chem. *275*, 19933–19941.

Harris, B.Z., and Lim, W.A. (2001). Mechanism and role of PDZ domains in signaling complex assembly. J. Cell Sci. *114*, 3219–3231.

Hillier, B.J., Christopherson, K.S., Prehoda, K.E., Bredt, D.S., and Lim, W.A. (1999). Unexpected modes of PDZ domain scaffolding revealed by structure of nNOS-syntrophin complex. Science *284*, 812–815.

Hung, A.Y., and Sheng, M. (2002). PDZ domains: structural modules for protein complex assembly. J. Biol. Chem. *277*, 5699–5702.

Im, Y.J., Park, S.H., Pho, S.H., Lee, J.H., Kang, G.B., Sheng, M., Kim, E., and Eom, S.H. (2003). Crystal structure of GRIP1 PDZ6-peptide complex reveals the structural basis for class II PDZ target recognition and PDZ domain-mediated multimerization. J. Biol. Chem. *278*, 8501–8507.

Iuliano, R., Trapasso, F., Sama, I., Le Pera, I., Martelli, M.L., Lembo, F., Santoro, M., Viglietto, G., Chiariotti, L., and Fusco, A. (2001). Rat protein tyrosine phosphatase η physically interacts with the PDZ domains of syntenin. FEBS Lett. *500*, 41–44.

Jannatipour, M., Dion, P., Khan, S., Jindal, H., Fan, X., Laganiere, J., Chishti, A.H., and Rouleau, G.A. (2001). Schwannomin isoform-1 interacts with syntenin via PDZ domains. J. Biol. Chem. *276*, 33093–33100.

Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard, M. (1991). Improved methods for binding protein models in electron density

maps and the location of errors in these models. Acta Crystallogr. A *47*, 110–119.

Kang, B.S., Cooper, D.R., Jelen, F., Devedjiev, Y., Derewenda, U., Dauter, Z., Otlewski, J., and Derewenda, Z.S. (2003). PDZ-tandem of human syntenin: crystal structure and functional properties. Structure *11*, 459–468.

Karthikeyan, S., Leung, T., Birrane, G., Webster, G., and Ladias, J.A. (2001a). Crystal structure of the PDZ1 domain of human Na^+/H^+ exchanger regulatory factor provides insights into the mechanism of carboxyl-terminal leucine recognition by class I PDZ domains. J. Mol. Biol. *308*, 963–973.

Karthikeyan, S., Leung, T., and Ladias, J.A. (2001b). Structural basis of the Na^+/H^+ exchanger regulatory factor PDZ1 interaction with the carboxyl-terminal region of the cystic fibrosis transmembrane conductance regulator. J. Biol. Chem. *276*, 19683–19686.

Karthikeyan, S., Leung, T., and Ladias, J.A. (2002). Structural determinants of the Na^+/H^+ exchanger regulatory factor interaction with the β_2 adrenergic and platelet-derived growth factor receptors. J. Biol. Chem. *277*, 18973–18978.

Koroll, M., Rathjen, F.G., and Volkmer, H. (2001). The neural cell recognition molecule neurofascin interacts with syntenin-1 but not with syntenin-2, both of which reveal self-associating activity. J. Biol. Chem. *276*, 10646–10654.

Kozlov, G., Banville, D., Gehring, K., and Ekiel, I. (2002). Solution structure of the PDZ2 domain from cytosolic human phosphatase hPTP1E complexed with a peptide reveals contribution of the β2-β3 loop to PDZ domain-ligand interactions. J. Mol. Biol. *320*, 813–820.

Kraulis, P.J. (1991). MOLSCRIP: a program to produce both detailed and schematic plots of protein structures. J. Appl. Crystallogr. *24*, 946–950.

Kurschner, C., Mermelstein, P.G., Holden, W.T., and Surmeier, D.J. (1998). CIPP, a novel multivalent PDZ domain protein, selectively interacts with Kir4.0 family members, NMDA receptor subunits, neurexins, and neuroligins. Mol. Cell. Neurosci. *11*, 161–172.

Laura, R.P., Witt, A.S., Held, H.A., Gerstner, R., Deshayes, K., Koehler, M.F., Kosik, K.S., Sidhu, S.S., and Lasky, L.A. (2002). The Erbin PDZ domain binds with high affinity and specificity to the carboxyl termini of delta-catenin and ARVCF. J. Biol. Chem. *277*, 12906–12914.

Lin, D., Gish, G.D., Songyang, Z., and Pawson, T. (1999). The carboxyl terminus of B class ephrins constitutes a PDZ domain binding motif. J. Biol. Chem. *274*, 3726–3733.

Maximov, A., Sudhof, T.C., and Bezprozvanny, I. (1999). Association of neuronal calcium channels with modular adaptor proteins. J. Biol. Chem. *274*, 24453–24456.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. D *53*, 240–255.

Navaza, J. (1994). AMoRe: an automated package for molecular replacement. Acta Crystallogr. A *50*, 157–163.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. In Methods in Enzymology, Volume 276, C.W. Carter, Jr., and R.M. Sweet, eds. (New York: Academic Press), pp. 307–326.

Schultz, J., Hoffmuller, U., Krause, G., Ashurst, J., Macias, M.J., Schmieder, P., Schneider-Mergener, J., and Oschkinat, H. (1998). Specific interactions between the syntrophin PDZ domain and voltage-gated sodium channels. Nat. Struct. Biol. *5*, 19–24.

Sheffield, P., Garrard, S., and Derewenda, Z. (1999). Overcoming expression and purification problems of RhoGDI using a family of "parallel" expression vectors. Protein Expr. Purif. *15*, 34–39.

Sheng, M., and Sala, C. (2001). PDZ domains and the organization of supramolecular complexes. Annu. Rev. Neurosci. *24*, 1–29.

Skelton, N.J., Koehler, M.F., Zobel, K., Wong, W.L., Yeh, S., Pisabarro, M.T., Yin, J.P., Lasky, L.A., and Sidhu, S.S. (2003). Origins of PDZ domain ligand specificity: structure determination and mutagenesis of the Erbin PDZ domain. J. Biol. Chem. *278*, 7645–7654.

Songyang, Z., Fanning, A.S., Fu, C., Xu, J., Marfatia, S.M., Chishti, A.H., Crompton, A., Chan, A.C., Anderson, J.M., and Cantley, L.C. (1997). Recognition of unique carboxyl-terminal motifs by distinct PDZ domains. Science *275*, 73–77.

Tochio, H., Zhang, Q., Mandal, P., Li, M., and Zhang, M. (1999). Solution structure of the extended neuronal nitric oxide synthase PDZ domain complexed with an associated peptide. Nat. Struct. Biol. *6*, 417–421.

Torres, R., Firestein, B.L., Dong, H., Staudinger, J., Olson, E.N., Huganir, R.L., Bredt, D.S., Gale, N.W., and Yancopoulos, G.D. (1998). PDZ proteins bind, cluster, and synaptically colocalize with Eph receptors and their ephrin ligands. Neuron *21*, 1453–1463.

Vaccaro, P., and Dente, L. (2002). PDZ domains: troubles in classification. FEBS Lett. *512*, 345–349.

Vaccaro, P., Brannetti, B., Montecchi-Palazzi, L., Philipp, S., Citterich, M.H., Cesareni, G., and Dente, L. (2001). Distinct binding specificity of the multiple PDZ domains of INADL, a human protein with homology to INAD from *Drosophila melanogaster*. J. Biol. Chem. *276*, 42122–42130.

Walma, T., Spronk, C.A., Tessari, M., Aelen, J., Schepens, J., Hendriks, W., and Vuister, G.W. (2002). Structure, dynamics and binding characteristics of the second PDZ domain of PTP-BL. J. Mol. Biol. *316*, 1101–1110.

Accession Numbers

Coordinates have been deposited with ID codes 1NTE, 1OBY, and 1OBX for the syntenin PDZ2 alone, PDZ2-syndecan-4 peptide complex, and PDZ2-IL5Rα peptide complex, respectively.